**Тестовое задание**

**Описание**

Вам представлена штатная численность предприятия, которое имеет офисы по всей России. Период наблюдения - 5 месяцев.

Сотрудник, отвечающий за подготовку отчетности на предприятии, при анализе таблицы столкнулся с некоторыми проблемами, которые мешают подготовить отчет руководству.

Помогите ему, изучите данные и решите поставленные задачи.

Представленные данные:

**Таблица Сотрудники:**

1) Период - формат Дата (пропуски не допускаются)

2) ID должности - целое число (пропуски не допускаются, уникальный на каждый период)

3) Табельный номер - целое число (допускаются пропуски)

4) Город - Текст (допускаются пропуски)

5) Должность - Текст (допускаются пропуски)

6) ФИО - Текст (допускаются пропуски)

**Таблица Города:**

1) Федеральный округ - Текст (пропуски не допускаются)

2) Регион - Текст (пропуски не допускаются)

3) Город - Текст (уникальный)

**Задачи**:

1. Предобработка таблицы:

1.1. В таблице есть ошибки данных (некорректный для поля тип данных, дубликат и пр.), укажите, как их обнаружили, и что предприняли для их устранения

1.2. Часть данных по некоторым должностям была утеряна, но известно, что ID имеет уникальную должность и локацию, и не меняется из периода в период, восстановите недостающие данные

1.3. Вытянуть строки с новыми штатными единицами, на которые еще не распределены должности и сотрудники. В графах Табельного номера и Должности проставить признак «Резерв»;

2. Построить сводную таблицу с разбивкой по месяцам со следующими показателями:

2.1. Количество действующих сотрудников (должностей);

2.2. Количество свободных должностей (подсказка: табельный номер пустой);

2.3. Доля свободных должностей от итого должностей;

3. Работа по дополнению предобработанной таблицы, добавьте следующие показатели:

3.1. Федеральные округа и Регионы;

3.2. Количество месяцев поиска сотрудника по должности накопительно по каждому месяцу (н-р должность "директор" не заполнена в январе до марта включительно, а в апреле заполнена, результат получится следующим: январь - 1, февраль - 2, март - 3). Задание повышенной сложности, если не получится реализовать, то напишите ваши идеи, как бы это могло быть реализовано.

4. Ответьте на следующие вопросы, построив сводные таблицы исходя из дополненных данных пункта 3;

4.1. В каких городах наблюдается нехватка кадров на последний месяц? Отсортируйте по убыванию;

4.2. В каком регионе работает наибольшее количество сотрудников на последний месяц? Отсортируйте по убыванию;

4.3. Какие должности подвержены наибольшей текучести (подсказка: расчет из п.2.3)? Выведите наименование должности и коэффициенты текучести по месяцам, отразите первыми те должности, по которым наблюдалась наибольшая текучесть.

5. Визуализация (тип диаграммы и библиотека на ваше усмотрение);

5.1. Количество действующих сотрудников по Федеральным округам;

5.2. Количество свободных должностей по регионам.

**Формат представления ответа:**

1) Рабочая книга в .ipynb (Jupyter Notebook)

2) Все ответы представлять там по пунктам, согласно тестовому заданию.